

OM of: US-09-805-550-2 to: GenBank: * out_format : pfs
Date: Aug 31, 2002 7:33 AM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODL=frim+g2n.model -DEV=xlh
-O=/cgn2_1/USPRO.spool/US09805550/tunat_29082002.160819-275/app-query.fasta_1.897
-DB=genbank1 -OPMT=fastap -SUFFIX=oligop2n.rge -GAPOP=4.500
-GAPEXT=0.050 -MINMATCH=0.100 -LOOPL=0.000 -LOOPEXT=0.000
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000
-DELOP=6.000 -DELEXT=7.000 -YGAPOP=60.000 -YGAPEXT=60.000
-STAR=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality
-THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTPMT=pfs -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09805550_ECGN1_1.4959 -NCPU=6 -ICPU=3 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-805-550-2
Query length: 405
Database: GenBank: *
Database sequences: 1797656
Database length: 187333701
Search time (sec): 3229.260000

WARN: XGAPOP and YGAPEXT must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

```
score_list:
Sequence      Strid Orig  ZScore  EScore Len  Documentation
gb.pat:AR153819 + 405.00 7208.10 0.0 1522 | AR153819 Sequence 1 from patent
gb.sts:G70483 + 38.00 662.12 1.6e-28 272 | G70483 721694431F073 maize leaf
gb.sts:G71250 + 29.00 501.84 1.4e-19 251 | G71250 721694431F073 maize leaf
gb.pl:AY037181 + 21.00 347.27 5.4e-11 1378 | AY037181 Arabidopsis thaliana
gb.pl:AY058196 + 21.00 347.27 5.4e-11 1426 | AY058196 Arabidopsis thaliana
gb.pl:LE8243875 + 20.00 329.07 5.8e-10 1498 | AJ243875 Lycopersicon esculent
gb.pl:AY063103 + 18.00 295.25 4.4e-08 1129 | AY063103 Arabidopsis thaliana
gb.pl:AY034912 + 18.00 293.25 5.7e-08 1458 | AY034912 Arabidopsis thaliana
gb.pl:DCRAD231 + 17.00 273.48 5.6e-07 1456 | Y12013 Daucus carota mRNA for
gb.pl:OS063530 + 14.00 221.71 0.0005 1533 | U63530 Oryza sativa cDNA RAD23 m
gb.pl:AB005248 + 14.00 194.94 0.0170 83646 | AB005248 Arabidopsis thaliana
gb.pl:AB005248 + 14.00 192.83 0.0223 114498 | AC006341 Arabidopsis thaliana
gb.pl:F309 + 13.00 204.14 0.0052 1466 | Y12014 Daucus carota mRNA for
gb.pl:DCRAD231 + 13.00 203.28 0.0058 1667 | AY038562 Arabidopsis thaliana
gb.pl:AY039562 + 13.00 203.14 0.0059 1702 | AR153820 Sequence 3 from patent
gb.pat:AR153820 + 13.00 178.17 0.1463 70957 | AP004631 Oryza sativa chromos
gb.hlg:AP004631 + 13.00 177.33 0.1629 80442 | AC010924 Arabidopsis thaliana
gb.pl:T24D18 + 13.00 176.57 0.1797 90149 | AC010924 genomic sequence for
gb.pl:AC010793 + 13.00 175.62 0.2029 110775 | AC021640 Arabidopsis thaliana
gb.hlg:AP003881 + 13.00 175.19 0.2144 110775 | AP003881 Oryza sativa chromo
gb.pat:AX314196 + 10.00 160.07 1.49 353 | AX314196 Sequence 7181 from pat
gb.pl:LE8270958 + 10.00 159.05 1.70 411 | AJ770958 Lycopersicon esculent
gb.ba:S76843 + 10.00 150.61 5.01 1450 | S76843 ori 3' of PGRS tandem r
gb.ba:AE006937 + 10.00 133.71 43.80 18101 | AE006937 Mycobacterium tuber
gb.ba:MY0035 + 10.00 129.28 77.31 31300 | AL001930 Mycobacterium tuber
gb.in:LMFL4738 + 10.00 127.95 91.79 35079 | AL135930 Leishmania major F1
gb.in:AF099922 + 10.00 125.30 128.85 42799 | AF099922 Caenorhabditis eleg
gb.hlg:AC027633 + 10.00 118.78 297.12 168283 | AC027633 Homo sapiens chro
gb.hlg:AC090643 + 10.00 118.71 299.93 170143 | AC090643 Homo sapiens chro
gb.hlg:AC026924 + 10.00 118.20 320.06 183514 | AC026924 Homo sapiens chro
gb.ba:AL646078 + 10.00 117.53 349.11 203050 | AL646078 Ralstonia solanace
gb.hlg:AC006880 + 10.00 116.29 409.10 244239 | AC006880 Caenorhabditis eleg
gb.hlg:AC006896 + 10.00 114.94 486.65 298960 | AC006896 Caenorhabditis eleg
gb.sts:G35919 + 9.00 142.03 15.07 362 | G35919 STS h14a1268.34 5, sequ
gb.sts:G36341 + 9.00 142.03 15.07 362 | G36341 STS h14a1268.5, sequ
gb.to:AB012865 + 9.00 137.93 25.49 668 | AB012865 Mus musculus mRNA for
```

```
seq_name: gb.pat:AR153819
seq_documentation_block:
LOCUS AR153819 1522 bp DNA linear PAT 08-ANG-2001
DEFINITION Sequence 1 from patent US 6235972.
ACCESSION AR153819
VERSION AR153819.1 GI:15121351
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1522)
AUTHORS Mahajan,P.B. and Taghian,L.
TITLE Maize Rad23 genes and uses thereof
JOURNAL Patent: US 6235972-A 1 22-MAY-2001;
FEATURES
source Location/Qualifiers
BASE COUNT 430 a 375 c 370 g 347 t
ORIGIN
alignment_scores:
Quality: 405.00 Length: 405
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-805-550-2 x AR153819 ..
Align seg 1/1 to: AR153819 from: 1 to: 1522
1 MetLysLeuAsnValLysThrLeuLysGlyThrAsnPhgIuIleGluAl 17
|||||
58 ATGAACTTACGTCAGACCCCTCAAGGCGACCACTTCGAGATCGAGGC 107
|||||
17 aSerProAspAlaSerValAlaAspValLysArgIleIleGluThrPhg 34
|||||
108 GAGCCCGATCGATCGGTGCTGATGGAAGATCATGATGAGACCTC 157
|||||
34 IngIyInSerThrTyArgAlaAspGlnGlnMetLysIleTyGlnGly 50
|||||
158 AAGGTCAAGATGACCTACCGGCGGACGCAAAATGCTCATATACCAAGG 207
|||||
51 LysIleLeuLysAspGluThrThrLeuGlnLysSerGlnGlyValAlaGluAs 67
|||||
208 AAAATTCAGAGATGAAACCACTTGGAAAGCAACGAGATGCTGAGAA 257
|||||
67 nSerPheLeuValIleMetLysSerLysAlaLysAlaSerSerSerGly 84
|||||
258 CAGCTCTCTTTATTAATGTTGTCACCAAGCTCAAGCATCATGAGTGAG 307
|||||
308 CTTCACCGCTACCTACCTCAAAAGCTCTGCACTGCGCCCAACCTGCT 357
|||||
101 AlaProValAlaProAlaAlaSerValAlaArgThrProThrGlnAlaPr 117
|||||
358 GCCCTGCGGCGCTGCTGATGATGCAAGATGCAAGCAACCAAGGCTCC 407
|||||
117 oValaIleThrAlaGluThrAlaProProSerValGlnProGlnAlaAla 134
|||||
408 TGTGTCACGACCTGAAAGCGACCTGCAAGTGTCCAACTGAGGCTGCTC 457
|||||
134 roAlaAlaIleThrValAlaAlaThrAspAlaAspValLysSerGlnAla 150
```

|||||
458 CAGCTGCTACGGTGTCTCTACTGATGATGCTGATGCTGACGAGCA 507
151 A1aserAnleuValPheglYasnAnleugInThrlleGInglInl 167
508 GCTTAAACCTGTGTATTGGCAACAATCTAGAACAGACTATCCAAAMAT 557
167 e1euaSpmetglYglYlYhTrpGluArGspThrValValArGAlAL 184
558 TCTTGACATGGGTGGTGTGTACATGGAAAGTGATCTGTTGCTGTGCTC 607
184 euArGAlALALeTyrAsnAsnProGluArGAlALeAspTyrleuYrSer 200
608 TACGTGCTGACATACATAACCCGAGAGAGCTATAGACTACTGTTATCT 657
201 G1YlEProGluAsnValGluAlaInProValAlArGAlALProAlAl 217
658 GGAATTCCTCGAATGTGAGGCTCAGCCTGTGGCCGAGCAGCTGCTGC 707
217 aGlyGInGlnThraSngInGlnAlaLaserProAlaGInProAlaValA 234
708 TGGCCAAACAACAAATCAGACAGCCGATCACCCGCTCAGCAGAGTTG 757
234 lAlaEProValGInProSerProAlaSerAlaGlyProAsnAlaAsnPro 250
758 CATTCACAGTCAGCCATCACCCTGCTGACAGGCTTAATGCAATCTCT 807
251 leuaSnleuPheProGInglYlYAlProSerGlyGlySerAsnProGlyVa 267
808 TTGAACCTTTTTCCTCAGGCTGTCCAAAGTGGTGGGTCCAAACCAAGTCT 857
267 lValProGlyAlaGlySerGlyAlaLeuAspAlaLeuArGInleuProG 284
858 TGTTCAGGTGCAGGATCTGGTCTCTGTATGCTTGTGTCACAGCTTCAC 907
284 lInPheGlnAlaLeuLeuGlnleuValGlnAlaAsnProGInleuGln 300
908 AATTTCAGACCTCTCTCAGTTAGTCAGGCTTAATCTCAAAATCTTGACG 957
301 PrometleuGlnleuGlnlyLySGlnAsnProGInleuArGleu1 317
958 CCAATGCTTCAAGAGCTAGGTAAACAAACCAAAATCTCGGTTGAT 1007
317 eGlnGlnAsnGlnAlaGluPheLeuArGleuValAsnGInSerProGluG 334
1008 TCAGGAAATCAAGCTGAGTTCTCGGCTTGGTGAATGATCTCTGAGG 1057
334 lYClYProGlyGlyAsnleuGlnleuGlnleuAlaAlaValProGIn 350
1058 GTGGTCTGGAGGAACTACTAGGTCACTGCGACGCTGTGTGCCACAA 1107
351 ThrLeuThrValThrProGluGluArGluAla1leGInArGleuGlu1 367
1108 ACCGTGACAGCTTACCCGAGAGGAGGAGGCTATCCACAGGCTGAGAGG 1157
367 yMetGlyPheAsnArGluLeuValleuGluValPhePheAlaCyAsnL 384
1158 AATGGGGTTCAAACCGTGAAGCTGTGTGTAAGTTTCTTGTGATGCAACA 1207
384 yAspGluGluLeuThrAlaAsnTyrleuLeuAspHisGlyHisGluPhe 400
1208 AGCAGAGAGAGCTTACAGCCCACTACCTCTGTGATCATGCGCATGATT 1257
401 AspAspGlnGlnGln 405
1258 GACGATCAGCAGCAA 1272
seq_name: gb_sts:G70483
seq_documentation_block: 272 bp DNA linear STS 08-JUN-2001
LOCUS G70483
DEFINITION 721694431FB73 maize leaf DNA Zea mays STS genomic, sequence tagged
site.

ACCESSION G70483
VERSION G70483.1 GI:14332168
KEYWORDS STS.
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
AUTHORS Yang, Y.J., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.
TITLE 3' UTR sequences of maize genes
JOURNAL Unpublished
COMMENT
Contact: Schnable, P.S.
Schnable Laboratory
Iowa State University
G405 Agronomy Hall, Ames, IA 50011, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schnable@iastate.edu
Primer A: TGTGCCCAACAGCTGAC
Primer B: TGTACTTAGAAAGCATTCGAC
PCR Profile:
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 45 seconds
Polymerization: 72 degrees C for 90 seconds
PCR cycles: 31
Thermal cycler: Perkin Elmer TC
Protocol:
Template: 10-20 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total vol: 20 ul
Buffer:
MgCl2: 2 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.4.
FEATURES
source Location/Qualifiers
1..272
/organism="Zea mays"
/strain="DB811"
/db_xref="taxon:4577"
/clone_id="maize leaf DNA"
/note="PCR products amplified from genomic DNA"
STS
BASE COUNT 64 a 59 c 78 g 71 t
ORIGIN
alignment_scores:
Quality: 38.00 Length: 38
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-805-550-2 x G70483 ..
Align seg 1/1 to: G70483 from: 1 to: 272
364 ArgLeuGlnGlyMetGlyPheAsnArGluLeuValleuGluValPhePh 380
|||||
13 CGGCTCGAGGGAATGGGTTCAACCGTGAAGCTTGTCTGAAGTTTCTT 62
380 eAlaCyAsnlyAspGluGluLeuThrAlaAsnTyrleuLeuAspHisG 397
|||||
63 TGCATGCAACAGGAGAGAGCTTACAGCAACTACCTGTGATCATG 112
397 lYHisGluPheAsp 401
|||||
113 GCCATGAGTTTGAC 126

seq_name:	gb_sts:c71250
seq_documentation_block:	
LOCUS	71250 251 bp DNA linear STS 08-JUN-2001
DEFINITION	721694.431FM017 maize leaf DNA Zea mays STS genomic, sequence tagged
ACCESSION	71250
VERSION	G71250.1 GI:14332935
KEYWORDS	STS.
SOURCE	Zea mays.
ORGANISM	Zea mays.
REFERENCE	1 (bases 1 to 251)
AUTHORS	Yang, Y.-J., Guo, L., Ashlock, D.A., Wen, T.-J. and Schnable, P.S.
TITLE	3' UTR sequences of maize genes
JOURNAL	Unpublished
COMMENT	

```

/translation="MKIFVKTLSSGSEFEIYKPADKVSQVTAETVYKAEYPAKAK
LIHQKVKLDETTLENNVENSFIVIMLSTKASPSGASTASAPAPSPOTVATP
QVSAPTASVPVTSATATAAPATAVOTVYGAASNLVAGTLESTVOOILDMG
GSMDDIVVRLRAAFNNPRAVEVLYSGIPAOEIPVPAQAPATGEALAMQANPOO
AAAPRAATGGRANPLNFPQMPADAGAGNIDFLRNSQFOALAMQANPOOIL
OPMIOELKONPOLVRLIOHOADFLRLINEVEEENVMQLEAMQAVTVPEER
FAIERLEBGMGDRAMVLEVFACNKNELANLYLIDHHEPEDO"
3'UTR
BASE COUNT      366 a      332 c      296 g      384 t
ORIGIN

alignment_scores:
      Quality:   21.00      Length:   21
      Ratio:     1.000      Gaps:     0
Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-805-550-2 x AY037181 ..
Align seg 1/1 to: AY037181 from: 1 to: 1378

292 ValGlnAlaAsnProGlnIleLeuGlnProMetLeuGlnIleuGly 308
|||||
875 GTACAAAGCAAAACCCACAAATTCTACAGCCTTACAGAGCTCGGTAA 924
308 sGlnAsnProGln 312
|||||
925 ACAAAACCCCCAG 937

seq_name: gb_p1:AY058196

seq_documentation_block:
LOCUS      AY058196      1426 bp      mRNA      linear      PLN 04-NOV-2001
DEFINITION Arabidopsis thaliana unknown protein mRNA, complete cds.
ACCESSION  AY058196
VERSION    AY058196.1 GI:16648837
KEYWORDS   FLI_CDNA.
SOURCE     thale cress.
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
            1 (bases 1 to 1426)
REFERENCE  1
AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
            Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
            Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
            Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
            Tortum,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
            Davis,R.W., Theologis,A. and Ecker,J.R.
            Arabidopsis cDNA clones
            unpublished
            2 (bases 1 to 1426)
TITLE     Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
JOURNAL   Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
REFERENCE Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
AUTHORS   Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
            Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
            Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
            Tortum,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
            Davis,R.W., Theologis,A. and Ecker,J.R.
            Direct Submission
            Submitted (28-AUG-2001) Salk Institute genomic Analysis Laboratory
            (SIGMAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT   RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA: "RIKEN
            Arabidopsis Full-length cDNA"): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

```

```

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,
Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J., Bowser,L.,
Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
Tang,C.C., Tortum,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
Davis,R.W., Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES
      source
      Location/Qualifiers
          1..1426
            /organism="Arabidopsis thaliana"
            /db_xref="taxon:3702"
            /chromosome="5"
            /clone="RAFL09-10-108(R1404)"
            /note="ecotype: Columbia"
          1..78
            /codon_start=1
            /product="unknown protein"
            /protein_id="AAL25609.1"
            /db_xref="GI:16648838"
            /translation="MKIFVKTLSSGSEFEIYKPADKVSQVTAETVYKAEYPAKAK
            LIHQKVKLDETTLENNVENSFIVIMLSTKASPSGASTASAPAPSPOTVATP
            QVSAPTASVPVTSATATAAPATAVOTVYGAASNLVAGTLESTVOOILDMG
            GSMDDIVVRLRAAFNNPRAVEVLYSGIPAOEIPVPAQAPATGEALAMQANPOO
            AAAPRAATGGRANPLNFPQMPADAGAGNIDFLRNSQFOALAMQANPOOIL
            OPMIOELKONPOLVRLIOHOADFLRLINEVEEENVMQLEAMQAVTVPEER
            FAIERLEBGMGDRAMVLEVFACNKNELANLYLIDHHEPEDO"
          1216..1426
            /db_xref="taxon:3702"

5'UTR
CDS
      3'UTR
BASE COUNT      377 a      343 c      301 g      405 t
ORIGIN

alignment_scores:
      Quality:   21.00      Length:   21
      Ratio:     1.000      Gaps:     0
Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-805-550-2 x AY058196 ..
Align seg 1/1 to: AY058196 from: 1 to: 1426

292 ValGlnAlaAsnProGlnIleLeuGlnProMetLeuGlnIleuGly 308
|||||
883 GTACAAAGCAAAACCCACAAATTCTACAGCCTTACAGAGCTCGGTAA 932
308 sGlnAsnProGln 312
|||||
933 ACAAAACCCCCAG 945

seq_name: gb_p1:LES243875

seq_documentation_block:
LOCUS      LES243875      1498 bp      mRNA      linear      PLN 28-JUL-1999
DEFINITION Lycopersicon esculentum mRNA for RAD23 homolog.
ACCESSION  AJ243875
VERSION    AJ243875.1 GI:5640110
KEYWORDS   RAD23 gene; RAD23 protein.
SOURCE     tomato.
ORGANISM   Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
            1 (bases 1 to 1498)
REFERENCE  1
AUTHORS   Lemarie-Chamley,M., Petit,J., Raymond,P. and Chevalier,C.
            Analysis of gene expression during early tomato fruit development

```

JOURNAL by mRNA differential display
 REFERENCE Unpublished
 2 (bases 1 to 1498)
 AUTHORS Lemaire-Chamley, M.
 TITLE Direct Submission
 JOURNAL Submitted (23-JUN-1999) Lemaire-Chamley M., Unite de Physiologie
 Vegetale, INRA-Bordeaux, B.P. 81, 33883 Villenave d'ornon cedex,
 FRANCE

FEATURES
 source Location/Qualifiers
 1..1498

gene
 CDS
 /organism="Lycopersicon esculentum"
 /variety="Cerasiformae"
 /cultivar="West Virginia 106"
 /db_xref="taxon:4081"
 /clone="LeRad23-8"
 /tissue_type="fruit"
 /dev_stage="6-10 DAA"
 108..1277
 /gene="rad23"
 108..1277
 /gene="rad23"
 /function="involved in DNA excision repair"
 /codon_start=1
 /product="RAD23 protein"
 /protein_id="CAB51544.1"
 /db_xref="GI:5640111"
 /translation="MKIFVTLKGTHTFEIEVPEKEDSVADVKKNIESVQSGDYPPAAQ
 MLHIGKVKLDITLLEENKVAENSEFVIMLSKNKSTGTSISALSNTPAQDGSSTO
 QARORITTPQATPALPOSASASAPPAVPAAASVTVDYDAAASNVASNLSTVVO
 QILDMGSSSRDLYVRLRAAYNNPERAVDLYSGITPOTEPYPAAPAPAYATAP
 ASQAQINPAQDASQAVPSSGPNAPLDLPQGLTNGSNGAGNLDPLRNSPQQA
 LRANVQAPQILQDELQKPNLMRLIDHOPDLRLINEPEGEGSNVIGQFAGA
 IPQAVTVPEREAEIERLEAMGFDRALVELYFACNKNKEILANLYLDLHHEFDE"

BASE COUNT 416 a 334 c 340 g 408 t
 ORIGIN

alignment_scores:
 Quality: 20.00 Length: 20
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-805-550-2 x LES243875 ..
 Align seg 1/1 to: LES243875 from: 1 to: 1498

292 ValGlnAlaAsnProGlnIleLeuGlnProMetLeuGlnIleuEnglyly 308
 |||||||
 948 GTCCAGCCCATCCGACATATGCGACCCATGCTCCAGAGTTGGGTAA 997

308 sGlnAsnPro 311
 |||||||
 998 GCAAAATCCA 1007

seq_name: gb_p1:AY063103

seq_documentation_block:

LOCUS AY063103 1129 bp mRNA linear PLN 26-NOV-2001
 DEFINITION Arabidopsis thaliana putative RAD23 protein (F20B17.8) mRNA,
 complete cds.
 VERSION AY063103.1 GI:17104776
 KEYWORDS FLI CDNA.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1129)
 Yamada, K., Banh, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D.,
 Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Tortum, M., Wu, H.C.,
 Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H.,

Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A.,
 Karlin-Neumann, G., Kawai, J., Kim, C., Koeseema, E., Lam, B., Lin, J.,
 Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,
 Sakurai, T., Satou, M., Seki, M., Shinozaki, K., Shinozaki, A.,
 Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 Full length cDNA of gene F20B17.8 (GI:7715605)
 Unpublished
 2 (bases 1 to 1129)

JOURNAL
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 COMMENT
 Submitted (15-NOV-2001) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 The RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
 Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and
 sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada, K.,
 Banh, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M.,
 Onodera, C.S., Quach, H.L., Tang, C.C., Tortum, M., Wu, H.C.,
 Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R.,
 Karlin-Neumann, G., Kim, C., Koeseema, E., Lam, B., Lin, J.,
 Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,
 Sakurai, T., Satou, M., Seki, M., Shinozaki, K., Shinozaki, A.,
 Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally
 to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP
 /PGEC) contributed equally to this work as PIs.
 Location/Qualifiers
 1..1129

FEATURES
 source
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="U09913"
 /note="This clone is in PUNI 51.
 ecotype: Columbia"
 1..1129
 /gene="F20B17.8"
 1..1098
 /gene="F20B17.8"
 /note="contains two UBA PF00627 domains and is a member of
 the ubiquitin PF00240 family"
 /codon_start=1
 /evidence="experimental"
 /product="putative RAD23 protein"
 /protein_id="AAL34277.1"
 /db_xref="GI:17104777"

CDS

translation="MKITVTLKGSHEFLRVLSPTIMAVKNIEDSGKDNPCGOO
 LIHNGKVLKDETSTVENKVTGKGVHFLVLSKSGSGSAGSAYVPSYTSSTKRAA
 PSTOSSPYRPSPIPAQOPAPOTDPTYGAAATSVSGSSLEQVOQIMEGGSDPKR
 TYTRALRAATNNPERAVDLYSGITPOTPAVAVPAPQIAGSGAAVAPASGSPNS
 LDLFPGETVAAGSGDLGFLERLNNDOQQLTAVHNSPQILQPMLELQKQNPQL
 RLIOENQAEFLQIVNEPYEGSDGEGMPQPEQEMPHAINVPAQDEAIOREAMGFD
 RALVIEAFIACDNEELANLYLLENSGDPED"
 576
 /gene="F20B17.8"
 /note="compared to genomic sequence"
 /replace="g"
 1099..1129
 /gene="F20B17.8"

misc_difference

BASE COUNT 332 a 254 c 261 g 282 t
 ORIGIN

alignment_scores:

Quality: 18.00 Length: 18
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-805-550-2 x AY063103 ..

Align seg 1/1 to: AY063103 from: 1 to: 1129

295 AsnProGlnIleLeuGlnProMetLeuGlnGluLeuGlyLygGlnAsnPr 311
|||||
769 AACCCGCAAAATCTTCAGCCCATGCTTCAGAACTTGGAACCAAAACCC 818
311 oGln 312
|||||
819 TCAA 822

seq_name: gb_pl:AY034912

seq_documentation_block:

LOCUS AY034912 1518 bp mRNA linear PLN 10-JUN-2001
DEFINITION Arabidopsis thaliana putative RAD23 protein (F20B17.8) mRNA,
complete cds.

ACCESSION AY034912

VERSION AY034912.1 GI:14334441

KEYWORDS

FLI-CDNA.

SOURCE

Arabidopsis thaliana

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

AUTHORS

Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
Goldsmith,A.D., Lee,J.M., Quach,H.L., Tang,C., Tortum,M., Yu,G.,
Bower,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
Kim,C., Koeseema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M.,
Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M.,
Shinozaki,K., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and
Theologis,A.
Full length cDNA of gene F20B17.8 (GI:7715605)
Unpublished

TITLE
JOURNAL
REFERENCE
AUTHORS

2 (bases 1 to 1518)
Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
Goldsmith,A.D., Lee,J.M., Quach,H.L., Tang,C., Tortum,M., Yu,G.,
Bower,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
Kim,C., Koeseema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M.,
Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M.,
Shinozaki,K., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and
Theologis,A.
Direct Submission
Submitted (10-MAY-2001) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RFLP CDNA (RAFL CDNA : 'RIKEN
Arabidopsis Full-length CDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

COMMENT

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL CDNA: Yamada,K., Liu,S.X.,
Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D.,
Lee,J.M., Quach,H.L., Tang,C., Tortum,M., Yu,G., Bower,L.,
Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Koeseema,E.,
Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J.,
Shinozaki,K., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)
contributed equally to this work as PIs.

FEATURES

Location/Qualifiers

source

1. 1518
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL05-08-C09 (R09913)"
/note="this clone is in a modified pluescript vector
(lambda ZAP) as a XhoI/SstI insert.
ecotype: Columbia
1. 136
/gene="F20B17.8"
1. 1518
/gene="F20B17.8"
137. 1234
/gene="F20B17.8"
/note="contains two UBA PF00627 domains and is a member of
the ubiquitin PF00240 family"
/codon_start=1
/evidence="experimental
/product="putative RAD23 protein"
/protein_id="AAK59419.1"
/db_xref="GI:14334442"
/translation="MKLVKTKLKGSHFEIRVLPEDTIMAVKNIEDSGKDNYPGGO
LIHNGKLDKSTLVENKVTKEGFLVYMLSKSGSAGASVQPSATTSRPPA
PSTQSSPVASPIPAQOOPAAQDTYGOAASITVSGSISLQWQIMWGSGNDK
TVTRALRAVNNPERAVDYLSGIPQAEVAVPEQIAGSGAAPAPASGPNSSP
LDLPQETVAAGSGDGTLEFLRNNDQFQDLRTVHNSNQIILQPMLOEGKQNPOL
RLIQENQAEFLQVNEPEYEGSDGEGDMQPEQDMRAINVTPAQOBAIORLEAMGD
RALVTEAFACDRNEELAAVNLLENSGDFED"
712

5'UTR

gene

CDS

misc_difference

/gene="F20B17.8"
/note="compared to genomic sequence"
/replace="g"
1235. 1518
/gene="F20B17.8"
1502
misc_difference
/gene="F20B17.8"
/note="not present in genomic sequence"

BASE COUNT 421 a 326 c 538 g 433 t
ORIGIN

alignment_scores:

Quality: 18.00 Length: 18
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-805-550-2 x AY034912 ..

Align seg 1/1 to: AY034912 from: 1 to: 1518

295 AsnProGlnIleLeuGlnProMetLeuGlnGluLeuGlyLygGlnAsnPr 311
|||||
905 AACCCGCAAAATCTTCAGCCCATGCTTCAGAACTTGGAACCAAAACCC 954

311 oGln 312
|||||
955 TCAA 958

seq_name: gb_pl:DCRAD23I

seq_documentation_block:

LOCUS DCRAD23I 1496 bp mRNA linear PLN 28-JAN-1999
DEFINITION Daucus carota mRNA for RAD23 protein, isoform I.
ACCESSION Y12013
VERSION Y12013.1 GI:1914682

KEYWORDS isoform I; RAD23 (ubiquitin-like protein involved in UV excision
repair).
SOURCE carrot.
ORGANISM Daucus carota

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; eusteroideae II; Apiales; Apiaceae; Daucus.

REFERENCE 1 (bases 1 to 1496)
 AUTHORS Sturm, A. and Lienhard, S.
 TITLE Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast
 JOURNAL Plant J. 13 (6), 815-821 (1998)
 MEDLINE 98345997
 REFERENCE 2 (bases 1 to 1496)
 AUTHORS Sturm, A.
 TITLE Direct Submission
 JOURNAL Submitted (24-MAR-1997) A. Sturm, Friedrich Miescher-Institut, Maulbeerallee 66, CH-4058 Basel, SWITZERLAND

FEATURES
 source 1..1496
 /organism="Daucus carota"
 /isolate="M001C"
 /sub_species="Queen Anne's Lace"
 /db_xref="taxon:4039"
 63..1211
 /function="assembly factor of the complex for nucleotide excision repair of V-damaged DNA"
 /codon_start=1
 /product="RAD23, isoform I"
 /protein_id="CAA72741.1"
 /db_xref="GI:1914683"
 /db_xref="SPTREMBL:003990"
 /translation="MKIYVKTKGQFEIQVNPDSVADVRSKRSIETAGAAVYPAQO MLITGKVKLKGDTLENNVAENSRIVIMLSKSPSGESTTSTAAPKAPQTSAP SVPAVAVSQPASTPVPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP ALOOILDMGGTMDIVYRIYRAAFNNPRAVEYLYSGIPROAPAPAPAPAPAPGOA ANPLDPPRAAPAPAPASGPNANPLDLPQCLPDMGSNAAGNIDPLRTNQPALR AVMSNPQIOLPMDLQKGNPHMLRLQEHQADPLQILNEMEGENLGHGPOAIS VTPERDAIERLEMGFDFRELVEYFEFCNNNELAANYLLDHMERFE"

BASE COUNT 403 a 361 c 331 g 401 t

ORIGIN

alignment_scores:
 Quality: 17.00 Length: 17
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-805-550-2 x DCRAD23I ..
 Align seg 1/1 to: DCRAD23I from: 1 to: 1496

295 Aaaprglrlleuglnpmetleuglnuengllysglnaap 311
 ||||||||||||||||||||||||||||||||||||||||
 906 AACCCCAAAATCTTGACAGCCCTTCAGAGATGGGAAACAAATCC 955
 311 o 311
 |
 956 T 956

seq_name: gp-pl:OSU63530

seq_documentation_block:
 LOCUS OSU63530 1533 bp mRNA linear PLN 07-NOV-1997
 DEFINITION Oryza sativa osRAD23 mRNA, complete cds.
 ACCESSION U63530
 VERSION U63530.1 GI:1488296
 KEYWORDS
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Euphorbiaceae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 1533)
 AUTHORS Schultz, T.F. and Quatrano, R.S.
 TITLE Characterization and expression of a rice RAD23 gene
 JOURNAL Plant Mol. Biol. 34 (3), 557-562 (1997)
 MEDLINE 97369378
 REFERENCE 2 (bases 1 to 1533)

AUTHORS Schultz, T.F. and Quatrano, R.S.
 TITLE Direct Submission
 JOURNAL Submitted (10-Jul-1996) Biology, University of North Carolina at Chapel Hill, CB 3280, Coker Hall, Chapel Hill, NC 27599-3280, USA

FEATURES
 source 1..1533
 /organism="Oryza sativa"
 /strain="Nipponbare"
 /db_xref="taxon:4530"
 58..1236
 /note="RAD23 homolog"
 /codon_start=1
 /product="osRAD23"
 /protein_id="AA065841.1"
 /db_xref="GI:1488297"

CDS

BASE COUNT 414 a 395 c 333 g 391 t

ORIGIN

alignment_scores:
 Quality: 14.00 Length: 14
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-805-550-2 x OSU63530 ..
 Align seg 1/1 to: OSU63530 from: 1 to: 1533

23 ValAlaaspvallysArgilellegluThrhrnglnlgln 36
 ||||||||||||||||||||||||||||||||||||||||
 124 GTGCTGATGTGAAAGATATCGAGCTACTCGACAGACG 165

seq_name: gp-pl:AB005248

seq_documentation_block:
 LOCUS AB005248 83646 bp DNA linear PLN 27-DEC-2000
 DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MXII0.
 ACCESSION AB005248 BM000015
 VERSION AB005248.1 GI:2264320
 KEYWORDS
 SOURCE Arabidopsis thaliana (strain:Columbia) DNA, clone:lib:Mitsui P1 clone:MXII0.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustroids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (sites)
 Sato, S., Kotani, H., Nakamura, Y., Kaneko, T., Asamizu, E., Fukami, M., Miyajima, N. and Tabata, S.
 TITLE Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence features of the 1.6 Mb regions covered by twenty physically assigned P1 clones
 JOURNAL DNA Res. 4 (3), 215-230 (1997)
 MEDLINE 97471969
 REFERENCE 2 (bases 1 to 83646)
 AUTHORS Nakamura, Y.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUL-1997) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research, 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: ynakam@kazusa.or.jp, Tel: 81-438-52-3935, Fax: 81-438-52-3934)
 COMMENT Address for correspondence: kas@kazusa.or.jp
 For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c-MXII0

Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Grail (informatics Group, Oak Ridge National Laboratory, <http://compbio.ornl.gov/Grail-1.3/>), GENSCAN (Chris Burge, MIT, <http://CCR-081.mit.edu/GENSCAN.html>), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.cbs.dtu.dk/services/NetGene2/>) and SplicePredictor (Volker Brendel, Stanford University, <http://gremli.ni.zool.iastate.edu/cgi-bin/sp.cgi>). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is MS17 and the 3' clone is MB18.

FEATURES

source

	<p>1. 83646 /organism="Arabidopsis thaliana" /strain="Columbia" /db_xref="taxon:3702" /chromosome="5" /clone="MX110" join(1791..2912,3016..3318,3584..4660) /note="gene_id:MX110.1" /codon_start=1 /evidence=not_experimental /product="disease resistance protein-1like" /protein_id="BAB09346.1" /db_xref="GI:9758812" /translation="MTEKIAIDVSNMINTSPSRDDELIGDAHMEKESLCLDS EVRMIGWSPSGIGKTTIARVLSQSENFELSTFMNIKELMTTRVCSDESAKIQ LQKQFQSOIINRDKMELHGVADRLNDKRVLLVDSIDQSLDAIAKTRFEGS SRILITFQDOKLKHGHNHIVKEEPAEAYOMFCMAFGONFPGDEELAEVY KLGLPLGRVMSHGRMSRHEWVALPRKIRLDASQSLIKESYDACEEDKL ELIACLPNNQMEVEEDVALSFIDVROGPHILAEKSLNKLKTSNCRIRBNHIL VQAGDIVRRKPGHQSICEPKKQFLIDAKDICEVLDNNGNRVAGIFLEVRNLSCQ LNISERAFDMNSKLFRLFDPDDESDKLYLPQGLNNLPQKLLIEMSRPMLCLS NFCRTYLVLRMNSKIQNMQNGQPGNLKRDLSKHLKELPDLSTANLEYLLM SGCISLVESSIGKLRKLMLSLRGCSKLEALPTNINLESIDYLDLCLIKKPE ISTINIKDLKLTAKIEKPESTIKSMHLKLEMSYSENLEKLPALDITITVYNDE MORLPQWVKKISHQTLGEGCKRLVYIPOLSDLSQLYVNCESLERLNSPQNH REIWEFLNCFKLNNEAREFTOTSTTHALPREPAPANTYANGSSIVNNHRLSTL LRRKACVLVKKIDNDEKAADRTTYIPRIRENDKIGVDPWRYRPHVPTILEDHL TFQIEADVTSNKLFESEIDKNNQAVIKEGVQL" complement(join(5504..5620,5688..5736,5836..5905, 5990..6122,6206..6250,6374..6688)) /note="gene_id:MX110.2 pit 1703994 similar to unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB09347.1" /db_xref="GI:9758813" /translation="MAASSLPSPFPCISYRKMPALITNSPRTTIPYLNKKGTQLIH KLPSGLKMEVIEQRKSKSERENPLPVFVHGSYHAAMWAMNMLPEFSSSGFDSVAVL KLGESDEPLGTAGTQTHASDIADIEBSNLGSSPVLVSGFGGLIVYLYANIN KRSLGTENAPPELGAVMCVSPSGNSGLVRLYLFSGPVAAFKVTLSLAKGQRKSI PLCRETFEQAMDOLYKRYNLAR" join(7788..7835,8043..8777,9032..9157,9205..9901, 10083..12938) /note="gene_id:MX110.3" /codon_start=1 /pseudo /evidence=not_experimental /product="non-LTR retroelement reverse transcriptase-1like protein" complement(13525..13827) /note="gene_id:MX110.4 pit 1702618 similar to unknown protein"</p>	
CDS	<p>/codon_start=1 /evidence=not_experimental /protein_id="BAB09348.1" /db_xref="GI:9758814" /translation="MKDVIIVYSLIKMTKLRSCPLDCCNSSETYFOLDKHLKNEHD LVLMVPOSSOWRCEOKMRHREYVDLMSATFPIPIAVVHHLIDTLIFRCLS" join(18792..18841,19220..19412,19494..19579,19675..19959, 20464..20607,20697..20783,20894..21006,21097..21162, 21216..21308,21350..21509) /note="gene_id:MX110.5 unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB09349.1" /db_xref="GI:9758815" /translation="TMSPKQISDGRSSHRNRPQIHLIGNFRIMSYMYRYLN QTGAPVLEFCCLPSSIFLIQKPKGKLSNOQIVPSLNGVITLAFYILMG LKSGPLRILISEYGAIVGLSVLGRGHVWKLPPTDSTETKEEYOTBOALCM MGMMIPYFACILSLRVIARVSLKNOQKRLHATITTSCTCLFVAMMDLIGSS SGKISELPFSMAFLSTIFGILIFVNDIAERLHMVSSPHLWAGCIIIVDI RVNMRKNISRTGEORLEVEIRNSIKRNPSPPIPYNDVLLVLDVNLVSGFIR FSIPSPSEFSEFCLEAVANILY" join(26073..30268,30416..31552) /note="gene_id:MX110.6" /codon_start=1 /pseudo /evidence=not_experimental /product="retroelement pol polyprotein-1like" 32202..33620 /note="contains similarity to retroelement pol polyprotein gene_id:MX110.7" /codon_start=1 /evidence=not_experimental /protein_id="BAB09350.1" /db_xref="GI:9758816" /translation="MVGCTRTNKRKSTMTSPQSEPVANHTREKKAIVTSP SPPRANRELEKGLLETTIDERTFMGDAFTSMDFPSETVFSEVTFMETETN PKIPLRAGELINFEVTKRAFSTIPDLCEAVGEFKSESPKRGTHFWMGELASG FSCNASIKSVNHPVIRYALKLAHLRGESTSTVSEMGFLQGVKELVEEDASG GDGLVHDDNDVNYACIPASVIEDYKRNLIKRTKSIYIGTLPLFPVAGVDSL FNAIPKREYIDENLVRSGLKRDQDAFYLYTRBNKLECLIPSKETINVESHE MFLPSKRYQDNLQEFMEIEEDPPSPRDDPVLANCDLPETFPVKTMEKFLQV IKSRKINKKSTVIGLKKVKKLQPPDYVSSDELYFPSSNEEDNDLDESNE TPATSSSRKKNRSRDSYTP" join(36310..37248,37409..37549,37633..37824,38363..38519) /note="emb CAB86928.1 gene_id:MX110.8 similar to unknown protein" /codon_start=1 /pseudo /evidence=not_experimental join(39564..40460,40627..40767,40853..41185) /note="gb AA055600.1 gene_id:MX110.9 strong similarity to unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB09351.1" /db_xref="GI:9758817" /translation="MDLSNLPYELCHILSTTRKAATLSVLSKRWRLIAFVNV DIYDHLRELPMDVDRVLAQGNPINKSLSDSGVSDRVDQWIDQVNVAGVSEIN LSIFVDSVFEDDYNLFPRFENKRLVGLSYISWLDGSLFLPLKLVLESYLSTVE KFETILHALPALLEIVNNNTYWKREDVDIVSESLKTLINFTVCTHLSLSDPL AYLSYSGVDDVAEAKENLFEARISLLVVEGISRVALINNDLLEDEYVLOE NVKMLNGINICITVLSPTNLEYSICSGSMVFKRLKSLSTKSAENRQWAPVLL RNCRPHLETIVLEGILHVTYPRKCGDACCVSREDKRGSLTSCPAKYLETIKFOGTTKX NMKHFLETFPCLKEKLMYMEENDPAQLRVPEVITGEMERRYKSSSCNQLLVGKX YKRWTP" join(42817..43766,43839..43979,44083..44403) /note="emb CAB86928.1 gene_id:MX110.10 similar to unknown protein"</p>	


```

CDS
    /codon_start=1
    /pseudo
    /evidence=not_experimental
    join(45089..45677,46843..47158,47312..47452,47544..47871)
    /note="contains similarity to non-LTR retroelement reverse
    transcriptase
    gene_id:MX110.11"
    /codon_start=1
    /pseudo
    /evidence=not_experimental
    48978..49433
    /note="gene_id:MX110.12
    unknown protein"
    /codon_start=1
    /evidence=not_experimental
    /protein_id="BAB09352.1"
    /db_xref="GI:9758818"
    /translation="MHVDSMKRARSAYEKEKRRSSNGEETREOSKHAMTAKAR
    MNYKESMKRPPPKKKSLMILITSLMRKEDTODDNNNTTYVEKEKEKGVONIRE
    NAEMLSSKVEEPMHLYEEEMTFEAFGGMWRDVAVMSGKSMFN"
    complement(join(50633..50872,51043..51177,51268..51438))
    /note="gene_id:MX110.13
    sp|P10798"
    /codon_start=1
    /evidence=not_experimental

alignment_scores:
    Quality: 14.00      Length: 14
    Ratio: 1.000        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
    US-09-805-550-2 x AB005248 ..

Align seg 1/1 to: AB005248 from: 1 to: 83646

299 LeuGlnProMetLeuGlnGluLeuGlyLysGlnAsnProGln 312
|||||
80121 TTTCAGCCTATGCTTCAGAGCTCGTAAACAAACCCCGAG 80162

seq_name: gb_pl:F309

seq_documentation_block:
    LOCUS      F309      114498 bp      DNA      linear      PLN 02-JUN-1999
    DEFINITION  Arabidopsis thaliana chromosome 1 BAC F309 sequence, complete
    sequence.
    ACCESSION  AC006341
    VERSION    AC006341.2  GI:4887257
    KEYWORDS   HMG.
    SOURCE     Chale cress.
    ORGANISM   Arabidopsis thaliana
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
    1 (bases 1 to 114498)
    Vysotskaya,V.S., Schwartz,J., Yu,G., Tortum,M., Lenz,C., Liu,S.,
    Lee,J., Li,J., Kremetskaia,I., Liu,A., Luros,J., Gonzalez,A.,
    Alatali,H., Araujo,R., Chao,Q., Conn,L., Conway,A.B., Dunn,P.,
    Hansen,N., Hutzar,L., Kim,C., Palm,C., Rowley,D., Shinn,P.,
    Walker,M., Davis,R.W., Ecker,J.R., Federspiel,N.A. and Theologis,A.
    The sequence of BAC F309 from Arabidopsis thaliana chromosome 1
    Unpublished (1999)
    2 (bases 1 to 114498)
    Theologis,A.
    REFERENCE 1
    TITLE      Direct Submission
    JOURNAL    Submitted (11-JAN-1999) Plant Gene Expression Center, 800 Buchanan
    Street, Albany, CA 94710, USA
    3 (bases 1 to 114498)
    Theologis,A.
    REFERENCE 2
    TITLE      Direct Submission
    JOURNAL    Submitted (25-MAY-1999) Plant Gene Expression Center, 800 Buchanan
    Street, Albany, CA 94710, USA
    4 (bases 1 to 114498)
```

```

AUTHORS      Theologis,A.
TITLE        Direct Submission
JOURNAL      Submitted (02-JUN-1999) Plant Gene Expression Center, 800 Buchanan
              St., Albany, CA 94710, USA
COMMENT      On May 25, 1999 this sequence version replaced gi:4139327.
              The sequence of BAC F309 from Arabidopsis thaliana chromosome 1.
FEATURES
    source
    1..114498
    /organism="Arabidopsis thaliana"
    /cultivar="Columbia"
    /db_xref="taxon:3702"
    /chromosome="1"
    /clone="F309"
    246..901
    /gene="F309.1"
    join(246..275,353..462,608..692,785..901)
    /gene="F309.1"
    /note="Similar to gb|Y12014 RAD3 protein isoform II from
    Daucus carota. This gene is probably cut off. EST
    gb|AA651284 comes from this gene."
    /codon_start=1
    /evidence=not_experimental
    /protein_id="AAD34676.1"
    /db_xref="GI:4966345"
    /translation="MVNSNPQILQPMLOELGKONPOLRLIOBNAEFLQLNEPYEG
    SLGDVDFDQPDQEMPHSVNTPPEQESIERLEAMGFDAIVIEAFLSCDRNEELAN
    YLLEHSADPED"
    complement(1361..2444)
    /gene="F309.2"
    complement(join(1361..1387,1506..1625,1899..2444))
    /gene="F309.2"
    /note="ESTs gb|T04357 and gb|AA595092 come from this
    gene."
    /codon_start=1
    /evidence=not_experimental
    /protein_id="AAD34673.1"
    /db_xref="GI:4966342"
    /translation="MGLNSKAEVANSRKNAABAEOKRQTREREQYWRBAPKSKA
    YKREEEAEKKAETAKLEAKRLAEQEELEKALKPKDKANRYTVVPVTEAEI
    IRRREDOVALAKKEDSKKOTRMAGDEVEKMYVTNTNDDSLIEAHVTDEALAR
    ITVSDMLPYDRHPERKRLKASFAYEEVELPRKSEKPGLTHTQYKDLVKKMKSSDN
    PINQAAAAANAE"
    3662..3562
    /gene="F309.3"
    join(3662..3916,4107..4475,4552..4669,4755..4978,
    5053..5562)
    /gene="F309.3"
    /note="Is a member of Pf100481 Protein phosphatase 2C
    family."
    /codon_start=1
    /evidence=not_experimental
    /protein_id="AAD34674.1"
    /db_xref="GI:4966343"
    /translation="MGLCSKIDKTRKETGATSTATTVEROSSGRLLRRPDLVSGG
    EISEIQVVGRLVNGSSSEIATVQGGKGTQDMMLWENFCSBTVLCGVFEGH
    GPGHVVSKVRVDMLEPFTLSQIKTTSGETQSSKNGLSAPCVDEWCELOCEK
    DKRLPEEMTLPKRLALTKCOQDKELKMHPTINCSCSTTSYTVKQKDLVGVIG
    DRAVLATKDQNALVAVOULDIKDLDESSESRIRHRCGRVYALDDEEVANVWLPJN
    SSPSGAMARDEPCLKDYGLISVDIVNHTREDOYITLATTGQVMVLSKRYVD
    IVASASRPTARAYVDVAVRNRMLKTPPSKNDCAVYVCLFEDTSAGCTSEKTYN
    HSHSESTSVTITSKADNRKEASTETETVPVWEIKKEKTPESCRISSKRTTAAEC
    ISVKDEEMSALEGLRVNSLSIPRFSGELRSSSRMWL"
    complement(10192..11349)
    /gene="F309.4"
    complement(join(10192..10323,10518..10562,10642..10742,
    10834..11039,11135..11349))
    /gene="F309.4"
    /note="ESTs gb|F15498, gb|H37515, gb|T41906, gb|T22448,
    gb|W43356 and gb|T20739 come from this gene."
    /codon_start=1
    /evidence=not_experimental
    /protein_id="AAD34675.1"
    /db_xref="GI:4966344"
```

gene
CDS
/translation="MASSSDSWARAYNEALKLSEINGMISERSSSAVTGPDQRAS
AIRKTIIFGNKLDLSQSLAEIHGAPITSEKNNRRDNGNLSKANQANALNMSN
FANRSLILGPDIKRDSRYTGNOGIVGQVNRDEDEGLEQETVWSTKRLA
LAVSELDQTLRLDLDYHVDYDSRLRVRQKSLAVNMKNMRSGSCMSMLTSLAI
VGLAVVIMWLKVKM"
/gene="F309.5"
/join(11987..12044,12084..12224,12408..12954,13133..13706)
/note="Contains two Pf101344 Kelch motif domains."
/evidence-not-experimental
/protein_id="A03467.1"
/db_xref="GI:4966346"
/translation="MPVSVSLPCQNEFLSDFSVSIIVGVFSGKVINFEARFMC
SKASRSGSEDFEITISRVRGSLPDLALICLISHYGHVLECVSRGRDILVAC
DYCYKARNMGSMFLFVTERSKNOMVAYADPEADRNPLPRTAYODGSMRDLAF
CVSNCLIVIGGCYVAFSPFHOKPVYTKYMRDPKMKKVAASMRPTTHACVS
VSGKVIYAGRNLTSHKGTSAEYTDYADRMETLPMAPRQNDGSLSTRGCTHVS
DVGFAONSSEVFNPDMTSTVEDVWPFSAQFAVQVKNDRVYTDWGSLLK
TRDDEGVNVSVPVLPNHPRELAFCYGFALRNELIYVIGKVLKWEESGAGR
FDIVLFFVAVCBPLDRPLMWRKPMICPAGCSIIICVSLSESSPP"
/complement(14717..17027)
/gene="F309.6"
complement(join(14717..15924,15993..16097,16178..17027))
/gene="F309.6"
/note="Similar to gb|A012423 wall-associated kinase 2
from Arabidopsis thaliana."
/codon_start=1
/evidence-not-experimental
/protein_id="A03467.1"
/db_xref="GI:4966347"
/translation="MGVDRFLVVMILLRICEAASSTPEPLRNCSDHGVSVY
PFGIGKCYKMKWEIVKSSDQPTLLPRLRATSTNLGRDPEISVYKRYIOS
PLKHSGCPNDRGSSSLNLGSEPFISENNKFAVCNNKAEVNVGLQVGETTC
GNEIRSKYKANTSCVGYKCCOMTIPPLQLOVPAVTEKLEQNOVAFLOFETIS
GSLFTEPELMESEYITELMELDLSYMTSKYKCLFEDFSYCCSGCLLFEQNA
YIRGGCODIDECBPHLNKCGKRCVAVLGSYREKTPALISTSSGCLLFEQNA
LICKANKRRVAKQKRFQNGGLLQDOTSLHGSVNTKYFSSNDLEAVDFEFA
SRLGOGGQGVYKGLMEDGMIYAVKSKALKENLEFNEILLQINRNVKIL
GCCLEVPFLVYEFIPNNLFDLHNPSSDFPMSVMEVRLCIAEVDALSYLSAVS
IPYHRDSTNILLDEKRAKYSDEGISVAIDHLLTIVGTYIGYVADALYSI
HFTKSDVYSFEGVILLTEKRPVSLRQEVSLGAYLEARNRRIHEILLARK
EEDDBREVLAVAKLARCLSLNSEHRTKMDVTEIDRMOSKRGTSQANQGEHNAH
IQIAMPESKSLSSPNIIVENSFSIDTFLPHKTKQ"
18899..23154
/gene="F309.7"
/join(18899..21385,21458..21538,21719..21871,21997..22160,
22371..22418,22491..22613,22684..22771,22955..23154)
/gene="F309.7"
/note="Contains Pf100069 Eukaryotic protein kinase domain.
ESTs gb|H37741, gb|T43005 and gb|A1100340 come from this
gene."
/codon_start=1
/evidence-not-experimental
/protein_id="A03467.1"
/db_xref="GI:4966348"
/translation="MDNRNRPHPQOHAMEGVYNDVPOGFPPDQTLGSNANRNP
ADVPGIHYISIQGEERFLEFLDRVYSOGEFIAAGDINPYGNHAGSEGSYV
SRMSVNGIRYOERTNPVHEFGNKGHLSHSAPEALICODRSIGNHAGSSAS
LTKAVVIGSGFKILPRGDSKRLRYGCGTHIITIRKDSMELROKVALEYIRRV
VKYQLGEDDALVYSCDDDLNMEYEMERKGSQKLRFLFSYSDIDALILY
NKSDBSEFYVAVNDMDLGSNSNLTNLDDSSANLLELDVRNTEGINGVPSOL
TGIDPOOSMOYSEAPTSFAQYPOIPIINGEFOFOQAVPNATLQYASNPSSSV
HYPOSILPNSTLOYPOSISSSYGLYPOYGEFOFOQAVPNATLQYASNPSSSV
YHPGITOONAVPOVEEPIKPKETKYDYDEPEREILATNHNOPADQTEVNRBP
SVATVPSODAAHMLPRBDTRONTPYKPSYDAVATVTOVPVSGEDOLSTSGRC
LVRTDSNSLIDIDYPERLOPTRRYRSEIRPREQLEMLRKSQSDLSQFLMSRP
QASTGOEPKAEAGISHEDSHIVNDENISGVNVAENETLDRKTVSGGGLTEARNL
SHVTERSHDIPKQTSVGLIDINDFPDQFLSEIFAKALSDMPKADSVYODGAG
VSLVNHNDPKNMSYFRNLADEQSDVDVAYIDRTPEFSDMEDGGEIARHAYAPLT
ENRVDPMQKTESSEFPAVYENLRTSDCEDEKSETRNAGLIPVPGSIADYDSGLQ
IIMNDLEELKEIGSGTFGVHGNKMGSDVALKRIKSCFAGSSQERLITGEFWE

gene
CDS
complement(23491..25496)
/gene="F309.8"
complement(join(23491..23643,23824..23922,24023..24104,
24189..24392,24473..24795,24882..25496))
/gene="F309.8"
/note="Similar to gb|L13612 DEAD-box protein (dbp45A) from
Drosophila melanogaster and is a member of Pf100270
DEAD/DEAH box helicase family."
/codon_start=1
/evidence-not-experimental
/protein_id="A03468.1"
/db_xref="GI:4966350"
/translation="MEPPPEEGCITIMSKRKNPTVYNISQKIDSDQNTPOFEK
FTNPPSDITTSATNFBGLAEVAVEICELGKRKPTPVQTHCVKTLGRDVLGLA
QTGSGKTAFAFLPILHRAEDPYGVFVLPVLEPDLVQLAEQFALGSLNLCVY
VGGMDMLQTMSTLVSRLVITPTGRILVLENNPDVAPVPSRTKFLVLEADVLY
GODELRTIFQCLKRSQTLFSAFMTSNLOALLEHSSNKAYFEAYEGKTVDTLQ
QFIEDDAKELLYVHILISQMEDGIRSAMIFVSTCRQCSIMLELEENIAMS
LNSQSMLSALSKRSKVPILLATDVASRGDLIPYDLYVNTDIPRDPDYVIRVGR

Alignment_scores:
Quality: 14.00 Length: 14
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:
US-09-805-550-2 x F309 ..

Align seg 1/1 to: F309 from: 1 to: 114498

299 LeuGlnPrometleuGlnGluGlnGlnAsnProGln 312
|||||
347 TTRACAGCCTATGCTTAAGAGCTCGGAAAGACAGACCCCA 388
|||||

seq_name: gb_pl:DCRAD23II

seq_documentation_block:
LOCUS DCRAD23II 1466 bp mRNA linear PLN 28-JAN-1999
DEFINITION Daucus carota mRNA for RAD23 protein, isoform II.
ACCESSION Y12014.1
VERSION Y12014.1 GI:1914684
KEYWORDS isoform II; RAD23 (ubiquitin-like protein involved in UV excision repair).
SOURCE carrot.
ORGANISM Eukarya; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
REFERENCE 1 (bases 1 to 1466)
AUTHORS Sturm, A. and Lienhard, S.
TITLE Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast
JOURNAL Plant J. 13 (6), 815-821 (1998)
MEDLINE 98345997
REFERENCE 2 (bases 1 to 1466)
AUTHORS Sturm, A.
TITLE Direct Submision
JOURNAL Submitted (24-MAR-1997) A. Sturm, Friedrich Miescher-Institut, Maulbeerallee 66, CH-4058 Basel, SWITZERLAND
FEATURES
source
1..1466
/organism="Daucus carota"
/isolate="W01C"
/sub_species="Queen Anne's lace"
/db_xref="taxon:4039"
40..1179
/function="assembly factor of the complex for nucleotide excision repair of V-damaged DNA"

```

/codon_start=1
/product="RAD23 protein, isoform II"
/protein_id="CAA72742.1"
/db_xref="GI:1914685"
/db_xref="SPTREMBL:003991"
/translation="MKLVKTLKGSFHEIRAPNDPVMAIKKNIEDLQKDNVPCGQO
LIHNGKVLKDESTLAESKISFDGLVYWLKSGKMSGSSPKAASSASAPAPAPAY
APAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP
GGGMDNNVNSALRAAANNPRAVDYLSGIPEMAEAVPVSHRQGDQINGNNNIS
DGVGAAPAGAPNSLPLNMFQETLSVGLAGLSLEFLRNPNFQETLSMVORNPDI
LQPMLELGKONPOLRQIOEHHEEFLQINPEVSESGDMFQEPQDVPEITVTAA
DQEAIRLEAMFDRGLVLEAFACDRNELAVNLEENAGDGED"
BASE COUNT      405 a      311 c      326 g      424 t
ORIGIN

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-805-550-2 x DCRAD23II ..

Align seg 1/1 to: DCRAD23II from: 1 to: 1466

182 ArgAlaLeuArgAlaAlaATyrAsnAsnProGluArgAla 194
|||||
553 CGTGCACCTCGGGCTGCTTACACATCCAGACGAGC 591

seq_name: gb_pl:AY039562

seq_documentation_block:
LOCUS      AY039562                1667 bp      mRNA      linear      PLN 20-JUN-2001
DEFINITION      Arabidopsis thaliana AT3g02540/F16B3_17 mRNA, complete cds.
ACCESSION      AY039562
VERSION        AY039562.1      GI:14517453
KEYWORDS
SOURCE        F11-CDNA.
ORGANISM      Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1667)
Kim,C.J., Chen,H., Cheuk,R., Koesema,E., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Dale,J.M., Gibson,H.A., Goldsmith,A.D.,
Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A.,
Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X.,
Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yu.G., Yu.S., Shinozaki,K.,
Davies,R.W., Theologis,A. and Ecker,J.R.
Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 1667)
Kim,C.J., Chen,H., Cheuk,R., Koesema,E., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Dale,J.M., Gibson,H.A., Goldsmith,A.D.,
Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A.,
Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X.,
Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yu.G., Yu.S., Shinozaki,K.,
Davies,R.W., Theologis,A. and Ecker,J.R.
Direct Submission
Submitted (04-JUN-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
Riken Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

```

```

The Salk, Stanford, PEGC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Shin,P., Chen,H.,
Cheuk,R., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bower,L.,
Dale,J.M., Gibson,H.A., Goldsmith,A.D., Jiang,P.X., Jones,T.,
Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L.,
Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yu.G., Yu.S.,
Davies,R.W., Theologis,A., and Ecker,J.R.
Shin,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.
Location/Qualifiers
1..1667
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="3"
/clone="RAFL02-02-B08(R12782)"
/note="ecotype: Columbia"
1..130
/note="putative RAD23"
/codon_start=1
/product="AT3g02540/F16B3_17"
/protein_id="AAK62617.1"
/db_xref="GI:14517454"
/translation="MKIFVTKLGSFHEIRAPNDPVMAIKKNIEDLQKDNVPCGQO
LIHNGKVLKDESTLAESKISFDGLVYWLKSGKMSGSSPKAASSASAPAPAPAY
APAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP
GGGMDNNVNSALRAAANNPRAVDYLSGIPEMAEAVPVSHRQGDQINGNNNIS
DGVGAAPAGAPNSLPLNMFQETLSVGLAGLSLEFLRNPNFQETLSMVORNPDI
LQPMLELGKONPOLRQIOEHHEEFLQINPEVSESGDMFQEPQDVPEITVTAA
DQEAIRLEAMFDRGLVLEAFACDRNELAVNLEENAGDGED"
BASE COUNT      458 a      392 c      373 g      444 t
ORIGIN

3'UTR
BASE COUNT      458 a      392 c      373 g      444 t
ORIGIN

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-805-550-2 x AY039562 ..

Align seg 1/1 to: AY039562 from: 1 to: 1667

163 ThrIleGlnGlnIleLeuAspMetGlyGlyGlyThrTTP 175
|||||
698 ACCATTCAGCAGATCTTGACATGGGTGAGGAAACCTGG 736

seq_name: gb_pat:AR153820

seq_documentation_block:
LOCUS      AR153820                1702 bp      DNA      linear      PAT 08-AUG-2001
DEFINITION      Sequence 3 from patent US 6235972.
ACCESSION      AR153820
VERSION        AR153820.1      GI:15121352
KEYWORDS
SOURCE        Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 1702)
AUTHORS      Mahajan,P.B. and Tagliani,L.
TITLE        Maize Rad23 genes and uses thereof
JOURNAL      Patent: US 6235972-A 3 22-MAY-2001;
FEATURES
source
1..1702
/organism="unknown"
BASE COUNT      490 a      373 c      432 g      407 t

```

ORIGIN

alignment_scores:
 Quality: 13.00 Length: 13
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-805-550-2 x AR153820 ..

Align seg 1/1 to: AR153820 from: 1 to: 1702

182 ArgAlaLeuArgAlaAlaTyrAsnAsnProGluArgAla 194
 ||||||||||||||||||||||||||||||||||||
 604 AGGGCTCTCCGTGCCGCTTACAAACAACCCCGAACGTCGT 642